

This is a repository copy of *Do existing real-world data sources generate suitable evidence for the HTA of medical devices in Europe? Mapping and critical appraisal.*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/174077/>

Version: Published Version

---

**Article:**

Drummond, Mike [orcid.org/0000-0002-6126-0944](https://orcid.org/0000-0002-6126-0944), Pongiglione, Benedetta, Torbica, Aleksandra et al. (9 more authors) (2021) Do existing real-world data sources generate suitable evidence for the HTA of medical devices in Europe? Mapping and critical appraisal. International Journal of Technology Assessment in Health Care. ISSN 0266-4623

<https://doi.org/10.1017/S0266462321000301>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

## Assessment

**Cite this article:** Pongiglione B *et al* (2021). Do existing real-world data sources generate suitable evidence for the HTA of medical devices in Europe? Mapping and critical appraisal. *International Journal of Technology Assessment in Health Care* 37, e62, 1–8. <https://doi.org/10.1017/S0266462321000301>

Received: 19 July 2020

Revised: 19 March 2021

Accepted: 21 March 2021

### Key words:



Real-world data; Real-world evidence; Medical device; Health Technology Assessment; Economic evaluations

### Author for correspondence:

Benedetta Pongiglione,

E-mail: [benedetta.pongiglione@unibocconi.it](mailto:benedetta.pongiglione@unibocconi.it)

# Do existing real-world data sources generate suitable evidence for the HTA of medical devices in Europe? Mapping and critical appraisal

Benedetta Pongiglione<sup>1</sup> , Aleksandra Torbica<sup>1</sup>, Hedwig Blommestein<sup>2</sup>, Saskia de Groot<sup>2</sup>, Oriana Ciani<sup>1,3</sup>, Sarah Walker<sup>4</sup>, Florian Dams<sup>5</sup>, Rudolf Blankart<sup>5,6</sup>, Meilin Mollenkamp<sup>7</sup>, Sándor Kovács<sup>8</sup>, Rosanna Tarricone<sup>1</sup>  and Mike Drummond<sup>9</sup>

<sup>1</sup>Centre for Research on Health and Social Care Management (CERGAS), SDA Bocconi School of Management, Bocconi University, Milano, Italy; <sup>2</sup>Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, Netherlands; <sup>3</sup>College of Medicine and Health, Institute of Health Research, University of Exeter, Exeter, UK; <sup>4</sup>Exeter University, Exeter, UK; <sup>5</sup>KPM Center for Public Management, University of Bern, Bern, Switzerland; <sup>6</sup>sitem-insel AG, Swiss Institute for Translational and Entrepreneurial Medicine, Bern, Switzerland; <sup>7</sup>Hamburg Center for Health Economics, Universität Hamburg, Hamburg, Germany; <sup>8</sup>Syreon Research Institute, Budapest, Hungary and <sup>9</sup>Centre for Health Economics, University of York, York, UK

## Abstract

**Aim.** Technological and computational advancements offer new tools for the collection and analysis of real-world data (RWD). Considering the substantial effort and resources devoted to collecting RWD, a greater return would be achieved if real-world evidence (RWE) was effectively used to support Health Technology Assessment (HTA) and decision making on medical technologies. A useful question is: To what extent are RWD suitable for generating RWE?

**Methods.** We mapped existing RWD sources in Europe for three case studies: hip and knee arthroplasty, transcatheter aortic valve implantation (TAVI) and mitral valve repair (TMVR), and robotic surgery procedures. We provided a comprehensive assessment of their content and appropriateness for conducting the HTA of medical devices. The identification of RWD sources was performed combining a systematic search on PubMed with gray literature scoping, covering fifteen European countries.

**Results.** We identified seventy-one RWD sources on arthroplasties; ninety-five on TAVI and TMVR; and seventy-seven on robotic procedures. The number, content, and integrity of the sources varied dramatically across countries. Most sources included at least one health outcome (97.5%), with mortality and rehospitalization/reoperation the most common; 80% of sources included resource outcomes, with length of stay the most common, and comparators were available in almost 70% of sources.

**Conclusions.** RWD sources bear the potential for the HTA of medical devices. The main challenges are data accessibility, a lack of standardization of health and economic outcomes, and inadequate comparators. These findings are crucial to enabling the incorporation of RWD into decision making and represent a readily available tool for getting acquainted with existing information sources.

## Introduction

Over the last few decades, digital innovation has permitted the generation, collection, and storage of a large volume of health-related data that can be employed to track patients' health and to monitor health service delivery and technologies during all stages of the lifecycle. Observational or administrative data that provide information on the routine delivery of health care and the health status of the target population are defined real-world data (RWD) (1). RWD can be medical health records, registries, biobanks, administrative data, health surveys, observational studies, health insurance data, data generated from mobile applications, etc. (2). The increasing availability of RWD has generated much attention in assessing whether, and if so to what extent, they can be used to generate clinical evidence regarding the usage, and potential benefits or risks, of medical technologies. In the Health Technology Assessment (HTA) glossary ([htaglossary.net](http://htaglossary.net)), evidence derived from RWD analyses is defined as real-world evidence (RWE).

RWE is particularly important for medical devices because, for this class of technology, available clinical evidence is traditionally of a lower standard, at least when compared with drug technologies. Clinical trials are often impeded by problems, including difficulties in

the blinding process, learning curve issues, hospital processes that may influence outcomes, incremental innovation, and fast product modification, with high loss to follow-up (3). In the literature, several issues with RWE use have been discussed by different stakeholders, often without distinguishing between medical devices and pharmaceuticals. Typical concerns with RWD use include confounding biases and limitations from measurement errors, selection bias, time-related bias, reverse causality, etc. (4;5). If RWD sources prove to be of adequate quality (i.e., based on suitable data collection protocols), it is fundamental developing a statistical analysis plan (e.g., checking covariate balance after applying the chosen confounding adjustment strategy, checking statistical power, evaluating positive or negative control outcomes, detailing the selection of the study population, and specifying primary vs. secondary analyses) (6). Possible RWE uses acknowledged in the literature and by the HTA community include regulatory processes such as market authorization (in the USA), postmarket surveillance, payer coverage, and reimbursement (7). RWE may be used over the whole product lifecycle, from the development of a new health technology, through the market access phase and post launch. RWE is recognized as a promising source of information for market access and reimbursement and as a complement to clinical trial evidence for treatment pathways, resource use, long-term natural history, and cost-effectiveness (8). RWD have also been increasingly used, in addition to randomized controlled trials (RCTs), for cost-effectiveness analyses (9) and for reassessments, payer coverage decisions, and outcome-based contracting (1;10). In a document issued in 2017, the FDA reported that RWD may potentially be used on their own or together with other evidence for understanding medical device performance at different points in the product lifecycle (11).

However, any potential use of RWD is conditional on their quality, relevance, and reliability. In a recent Health Technology Assessment International (HTAi) Policy Forum, it was noted that there is a lack of agreement between different parties regarding what data are needed, when, and for what purpose. Indeed, there is no clear consensus among stakeholders about when to use RWD (1): “The HTA community is currently standing at a cross-road, as it is not yet fully equipped to address these key challenges” (12). Therefore, it is important to explore these issues, especially for health technologies such as medical devices, where HTA may rely heavily on RWD.

This study aims to offer new tools for addressing some of these challenges, providing empirical evidence on the use of RWD for the HTA of medical devices. To achieve this general goal, we (i) systematically mapped RWD sources in Europe for three selected case studies: hip and knee arthroplasty, percutaneous transcatheter valve replacement technology (transcatheter aortic valve implantation [TAVI] and transcatheter mitral valve repair [TMVR]), and procedures performed by the da Vinci Surgical System. Then, we (ii) provided a comprehensive assessment of their content and evaluated their appropriateness for conducting the HTA of medical devices.

## Methods

### Choice of Case Studies

RWD source mapping was performed for the three case studies of medical devices: (i) *hip and knee arthroplasty*, (ii) *TAVI and TMVR*, and (iii) *the da Vinci Surgical System*. The choice of

case studies was not intended to be representative of all types of medical devices, but rather to cover a spectrum of heterogeneous cases in terms of the epidemiology of diseases, demographic trends (e.g., population ageing), characteristics of the procedures, and the maturity and type of the technology. Supplementary Table 1 shows, in a comparative way, the key characteristics of the three medical technologies selected, providing details on their maturity, EU classification, indication, and epidemiological/demographic characteristics and forecasts.

### Types of RWD

The identification of RWD sources for each case study was based on an adaptation of the classification of Makady *et al.* (13). To get a complete overview of the main RWD sources, we considered the following sources and study designs for inclusion: (i) *administrative data*, (ii) *registry data*, and (iii) *other observational data*. Data that could not be assigned to any of the categories were defined as (iv) *other data*. Registries are a specific type of observational data, where information about the health status of patients and the health care they receive over varying periods of time are typically recorded. Given that they play a prominent role in market/clinical surveillance, we included them as a separate category, alongside other observational data, such as health surveys and hospital data.

### Search Strategy

To map RWD sources for medical devices, multiple search strategies were adopted and implemented for each case study. We performed a systematic literature review, a targeted search of the gray literature and the Web sites of governments, research institutes, and relevant public agencies. In addition, we sought advice from experts in the devices and procedures/diseases of interest.

The systematic search was performed using PubMed. The search was developed to combine terms referring to the selected RWD types, with either the disease or the procedure or the medical device pertaining to the case study, and with the countries included in the mapping. The set of key words is available in Supplementary Table 2. The inclusion and exclusion criteria were set so as to match RWD definitions, the diseases, procedures, and medical devices, and geographic settings. More specifically, studies that used data sources not listed among selected RWD types were excluded; studies that used selected RWD types, but that were conducted by a single research unit (e.g., one hospital and patients of one surgeon) were excluded because we aimed to map RWD sources that provided generalizable data. This “single-center” exclusion criterion was not applied to da Vinci robotic surgery because less evidence is available, and single research unit studies often represent the only available RWD source. Furthermore, single-center studies were included in the first and second case study when considered particularly relevant (e.g., because many patients were included). There were other exclusions: nonempirical studies (e.g., literature reviews and commentaries), studies dealing with neither the disease nor procedure nor device of interest, studies outside the countries of interest, and studies based on data collected before 2013 because we wanted to identify information relevant to present-day decision making. For each search, the screening and study selection were illustrated through an adapted version of the PRISMA flowchart.

Gray literature scoping consisted in screening national/European sources, such as webpages and online archives to gain a general overview of the accessibility and breadth of data

knowledge related to each case study. There was also a nonsystematic search on Google Scholar.

At least one author from the countries represented in the research team (namely, Italy, Germany, England and Wales, Netherlands, Hungary, and Switzerland) performed the search for his or her country, so that it was possible to consult relevant national sources in the national language. Additional European countries, not represented in the research team, were also included, and these were selected based on the authors' knowledge of local language and context (see Supplementary Table 3). International sources, involving multiple European countries, were also mapped and were given a separate setting.

Finally, after performing a systematic search and gray literature scoping, advice from experts with the devices, procedures, and/or diseases was sought. This was done to assess the completeness and quality of the mapping.

### Information Extraction

Information on each of the RWD sources was extracted and entered into a spreadsheet that synthesizes the general features of the sources and the variables included and that provides the references and links for each source. The extraction template was created for each case study, with each row providing information on a single RWD source. In Supplementary Tables 4 and 5, the general framework of the template and definitions for each field are presented.

### Results

In total, mapping covered fifteen countries, plus multinational/European sources. We identified seventy-one RWD sources on hip and knee arthroplasties, ninety-five sources on TAVI and TMVR, and seventy-one on the da Vinci surgical system. A complete list of sources with full details is provided in the extraction template available in Supplementary Table 6. Supplementary Figure 1 shows PRISMA flowchart with the screening results of the PubMed search.

The number of sources varied substantially across countries for each case study. Germany was the country with the highest number of sources, followed by multinational sources and Italy (Supplementary Figure 2). Of the seventy-one sources found for arthroplasty, almost half (thirty-four) are *registry data*, 32 percent (twenty-three) *other observational data*, and 17 percent (twelve) *administrative data*, with two sources categorized as *other data*. *Other data* are the Network of Orthopaedic Registries of Europe and European Arthroplasty Registry. These are international registry networks rather than actual registries, so they are categorized separately. A similar distribution of sources is observed for RWD on TAVI and TMVR: 60 percent are *registry data* (fifty-seven), 34.7 percent are *other observational data* (thirty-three), and 5.3 percent are *administrative data* (five). For the da Vinci robot, 84.5 percent are classified as *other observational data* (sixty), 11 percent are *registry data* (eight), and 4 percent *administrative data* (three).

### General Features

The general characteristics of the selected sources were the data aggregation level (i.e., the unit of analysis), data accessibility, geographic coverage, and the approach for selecting the sample. The results are synthesized in Table 1 by case study, and discussed

**Table 1.** General features of RWD source by case study, % where not stated otherwise

	Arthroplasty (N=71)	TAVI and TMVR (N=95)	da Vinci robot (N=71)
Aggregation level (%)			
Patient	88.7	94.7	94.4
Hospital	2.8	3.2	4.2
Region (NUTS2)	0	0	1.4
Country	8.5	2.1	0
Data accessibility (%)			
Public	7.1	1.1	0
Restricted	57.7	34.7	28.2
Private	16.9	36.8	53.5
Other	18.3	27.4	18.3
Geographical coverage (%)			
Single center	na	na	49.3
Subnational/multicenter	32.4	29.5	15.5
National	56.3	35.8	26.8
Transnational	11.3	34.7	8.5
Inclusion approach (%)			
Disease-based	45.1	29.5	28.2
Single device-based	15.5	32.6	38.0
Multidevice-based	21.1	33.7	14.1
Other/unknown	18.3	4.2	19.7

NUTS2, Nomenclature of Territorial Units for Statistics level 2.

here, distinguishing the RWD type. In terms of their aggregation level, most sources had, in all case studies, patient-level data, but the proportion was lower for arthroplasties. Accessibility to data in most cases was limited, either restricted or private (distinction of terms provided in Supplementary Table 5). Nonregistry observational data were the most difficult to access; hence, data on the da Vinci robot, that mainly came from this source, were most commonly private. Regarding geographical coverage, the highest proportion of transnational sources was for TAVI and TMVR, whereas national coverage was particularly common for arthroplasty, due to the national registries set up in most European countries. Single-center sources of RWD were included only for the da Vinci robot, and they represented half of the existing sources. Finally, for arthroplasty, around 45.1 percent of sources selected patients based on their disease and 36.6 percent based on medical device (either single or multiple devices). For TAVI and TMVR, two thirds of RWD sources were either single or multiple medical device-based. This mainly depended on the fact that two prostheses are most commonly used in clinical practice, the Edwards SAPIEN valve (Edwards Life-sciences, Irvine, CA, USA) and the CoreValve System (Medtronic Inc., Minneapolis, MN, USA). Many studies were sponsored by these manufacturers. For robotic surgery, almost half of the RWD sources were either single- or multiple-device-based. Many studies compared robotic surgery with laparoscopic or open interventions, and their inclusion approach was classified as "other."



## Relevance for HTA Purpose

To assess whether RWD sources can be used to conduct HTA, we asked whether they included information on health and economic outcomes and comparators.

## Health Outcomes

Almost all RWD sources included in our mapping included health outcomes that are relevant for HTA. For arthroplasty, four sources did not include data on health outcomes or information was not retrievable (see spreadsheet). All sources on TAVI and TMVR included at least one health outcome. For the da Vinci robot, three German observational studies did not include health outcomes or information was not available.

Some health outcomes were common to all case studies: mortality, readmission (for hip/knee prostheses this often corresponds with revision or reoperation), and patient-reported outcome measures (PROMs). PROMs were more commonly available for arthroplasty, including pain assessment, using the Western Ontario and McMaster Arthritis Center (WOMAC) score and the Knee Society 18 Score. Instruments for measuring health-related quality of life issues, such as EuroQol 5D (EQ5D), the Health Utilities Index (HUI), and the Short Form 36 health survey (SF-36), were also available for a minority of RWD sources.

For TAVI and TMVR, a consortium of experts, the “Valve Academic Research Consortium” (VARC), established clinical end points and standardized definitions (14). In 2011, the first consensus document was published, and 2 years later, the selection and definitions of end points were revised and updated, and named VARC-2 (15). In the selected sources, more than half of RWD included end points defined according to either VARC (twelve sources) or VARC-2 (forty-one sources).

The da Vinci case study included a broader variety of health outcomes, given that we did not restrict the use of robotic surgery to specific diagnoses. Health outcomes included intraoperative outcomes (forty-one sources), among which the most reported

was blood loss (twenty-four sources), conversion to open/other surgery (twenty-three sources), and postoperative complications and relapse/recurrence (forty-four sources). We classified a residual group of health outcomes as “other.” They included very heterogeneous measures, such as recovery to normal breathing, swallowing functions, and the oncological adequacy of resection.

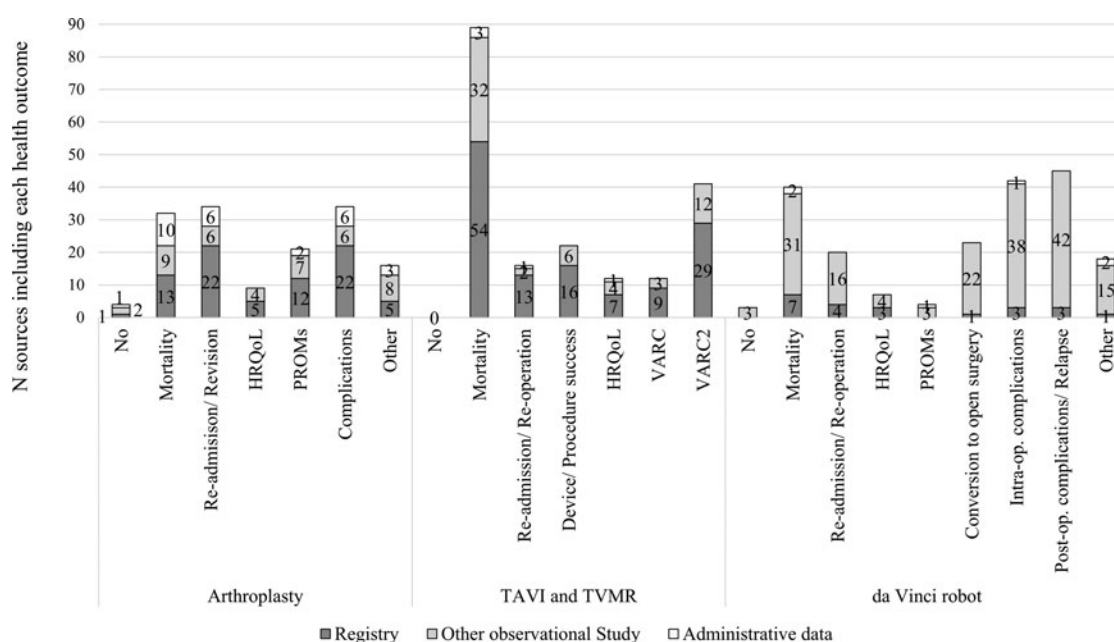
Figure 1 shows the health outcomes for each case study and by RWD type, selected based on how frequently they were available and their similarity across case studies.

## Resource Use

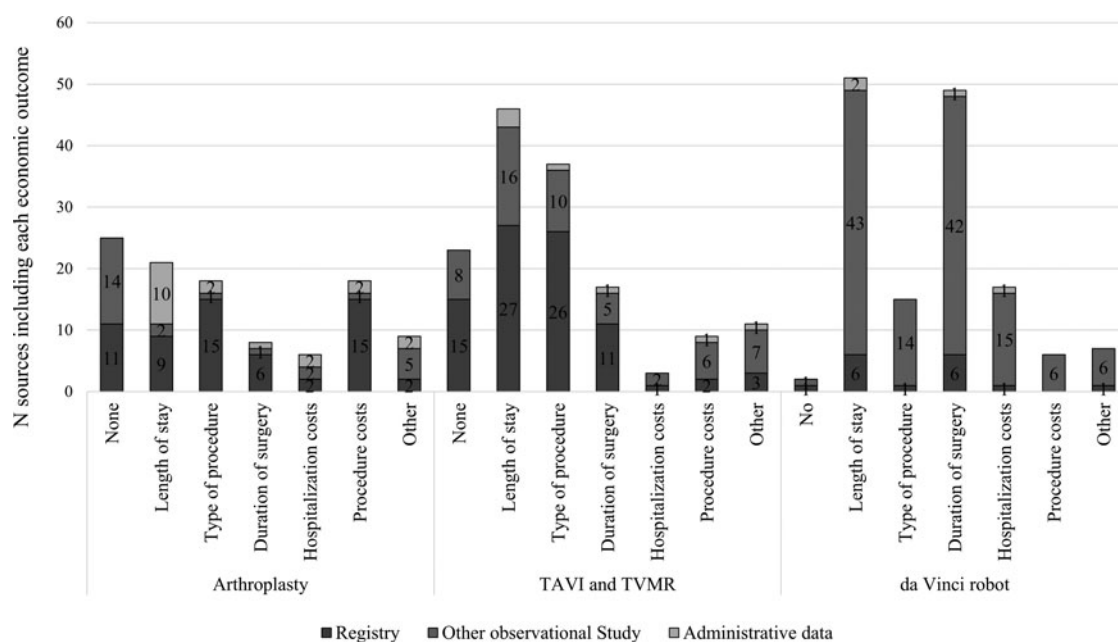
Compared with health outcomes, economic outcomes were less frequently available, especially in the case of arthroplasty, for which about one third of registries (eleven out of thirty-four) and more than half of observational studies (fourteen out of twenty-three) did not provide economic information.

For all case studies, the most commonly available information was length of stay, which included various types such as hospital and ICU length of stay, preintervention and postintervention length of stay and so on. The other common economic outcome was the type of procedure or procedure approach. In the case of arthroplasty, this referred to whether the procedure was a revision or replacement, access, and fixation methods. For TAVI and TMVR, it typically included whether the access route was transfemoral, transaortic, trans-subclavian, or transapical (16). The duration of surgery appeared particularly important for the da Vinci robot and was reported in almost 70% of sources (forty-nine). Hospitalization and procedure costs were composite and mixed categories, and were often not described in detail. The residual “other” group was composed of heterogeneous information. This included, for example, outpatient visits, employment status, and the reduction of earning capacity. These were costs relative to follow-up complications such as antithrombotic treatment, the use of aortography, or echocardiography.

Figure 2 shows the economic outcomes for each case study and by RWD type. These selections are based on how frequently they were available and how similar they were across case studies.



**Figure 1.** Number of RWD sources that include health outcomes, by type of health outcome and case study. *Note:* HRQoL, health-related quality of life; PROMs, patient-reported outcomes; VARC, Valve Academic Research Consortium.



**Figure 2.** Number of RWD sources that include economic outcomes, by type of resource and case study.

### Comparators

For arthroplasty, eighteen sources (almost 25%) did not include comparators suitable for HTA, and for six sources, the information was unknown. The most commonly available comparators were medical devices to be compared with other (types/versions of) medical devices. This was most commonly the case for registries, where hip and knee prostheses were best traced. In other cases, it was possible to compare the characteristics of different devices (e.g., materials and characteristics of the components).

For TAVI and TMVR, thirty-one sources did not include comparators, mostly registries (twenty-one) and almost no international sources allow to compare the medical device to other devices or clinical procedures. In some of these cases, comparisons were made in terms of route of access (transfemoral vs. others). When comparisons were possible, they were either between TAVI-specific devices (e.g., Edwards Sapien vs. Medtronic CoreValve)—this was typical in multidevice-based registries, or between TAVI procedures and other procedures, commonly surgery aortic valve replacement (SAVR).

For robotic surgery, twenty-six sources did not have comparators for assessing the da Vinci robot, and for ten sources, it was difficult to find such information. Almost half of the studies compared robotic procedures with other types of surgeries, typically laparoscopic or open surgeries or both. In some cases, comparisons were made between different versions of the da Vinci robot.

Figure 3 shows the comparators available in sources for each case study and by RWD type, selected based on how frequently they were available in each data source.

### Discussion

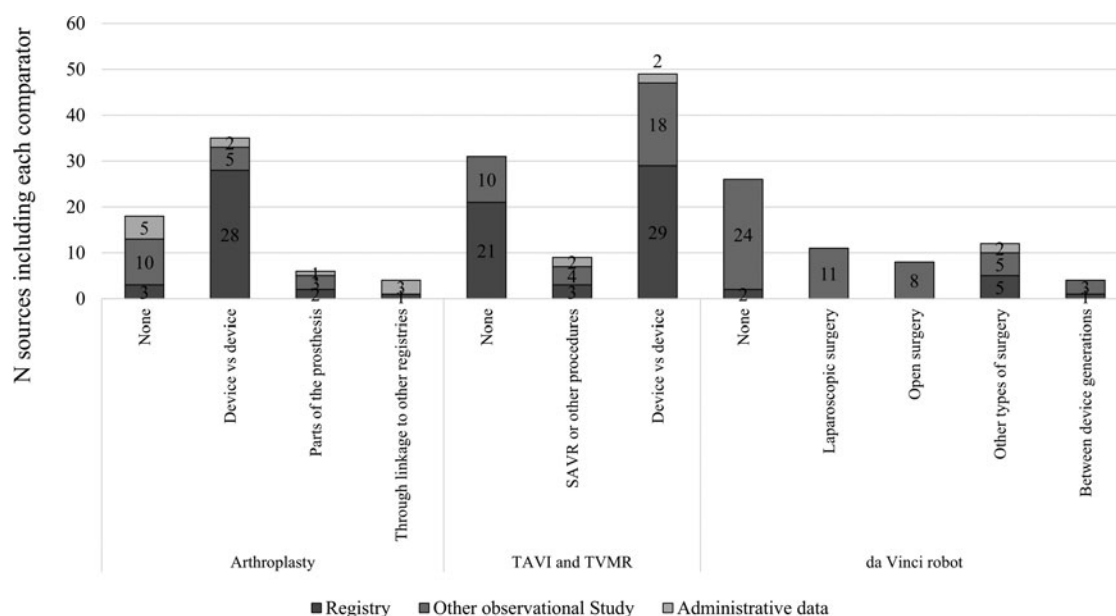
RWD collection for medical devices in Europe is extensive and growing, and the interest in using RWD to produce RWE has increased in parallel. The key RWE uses for medical devices include epidemiologic and safety evaluation, the characterization of treatment patterns, and healthcare utilization trends (17). The FDA has recently encouraged RWE use for regulatory

purposes (11). With the enactment of the new European Medical Device Regulation, RWE will likely become an important surveillance tool in Europe too. Given the considerable effort and resources devoted to the collection of RWD, a greater return would be achieved if the data were also useful for HTA. In this work, we sought to assess, through a selection of medical devices, whether existing sources are suitable for HTA and what, if anything, should be done to maximize their potential.

We mapped existing RWD sources in Europe and critically assessed whether they could be used for the HTA of medical devices. We found that, depending on the characteristics of the technology and the stage of the product lifecycle, certain types of RWD sources are more commonly available. For the most mature medical device—the endoprosthesis for arthroplasties—registries are the most important source of RWD. The other two technologies are, relative to the endoprosthesis, more recent, and single- or multicenter observational studies are the main sources of RWD here. This is especially true for the da Vinci robot, which is a complex and expensive technology. For the two technologies, most of the existing data sources are private, and cross-technology comparisons are lacking.

Our mapping depicted a heterogeneous scenario across countries. Germany stands out for the number of RWD sources; this has already been reported for cardiology (18). Some other countries seemed to lack RWD sources. This was particularly true for some Eastern European countries. International or multicountry studies proved a very important source, especially for TAVI and TMVR and for the da Vinci robot.

In terms of the suitability of existing sources for HTA purposes, several barriers emerged. Data accessibility is largely restricted; there is a lack of standardization of health and economic outcomes between countries and regions; economic outcomes are included in most sources but are rather generic and do not allow for the estimation of full costs; comparator(s) often do not allow for comparisons between technologies, and their data on health outcomes and resource use are not always available. Finally, and critically, data quality and completeness,



**Figure 3.** Number of RWD sources that include comparators, by type comparator and case study.

and the availability of demographic and clinical variables to control for confounding vary from data source to data source. Data integrity, meanwhile,—which refers to the accuracy and consistency of data collected over the lifecycle of a medical technology—proves difficult to assess.

The question of accessibility, or rather inaccessibility, has important implications. Academic and nonacademic research is strongly limited if existing data are not accessible, or if their existence is not disclosed. The finding that most RWD are *not* accessible is a call for policy makers and regulators to take action to facilitate data access, ensuring patients' privacy and data protection, even when data come from private stakeholders.

Regarding standardization, there is a clear need for consistent and appropriate selection, measurement, use, and reporting of outcomes in clinical research and practice. The inconsistencies and biases due to incomparable data on the effects of interventions could be addressed with the development and application of agreed standardized outcome sets, known as core outcome sets (19). An attempt in this direction is represented by the clinical end points for percutaneous transcatheter valve replacement, for which a common set of outcomes have been established, VARC, or VARC-2 end points. Another example is the joint collaboration and cross-country coordination promoted by the European Federation of National Associations of Orthopaedics and Traumatology that has created the Network of Orthopaedic Registries of Europe. This network supports the development of national and transnational arthroplasty registries and the development of a minimum arthroplasty data set to enhance the comparability of reports through standardization.

Overall, information on resource use is limited and includes general indicators, such as length of stay and information on the type of procedure. Only rarely specific costs related to the procedure type are included. Some variables, such as rehospitalization and reoperation, here, were presented as health outcomes, but they may also be considered as economic outcomes. Importantly, for some registries, mostly in Scandinavian countries, data linkage is possible and, therefore, even if registries do

not directly include economic outcomes, these can be obtained and/or expanded through other administrative data sets. It is very important to enable as much as possible data linkage (20), to avoid enlarging existing data sets with additional variables, which is expensive in terms of costs and time, and to capitalize on existing sources.

Comparators are often inadequate in comparative-effectiveness research. In the case of TAVI, for example, there are important ongoing discussions on its suitability for patients suffering from severe aortic stenosis at low and intermediate risk (21). The number of studies and data that allow only specific and limited comparisons (e.g., old TAVI vs. improved TAVI, between procedure approaches, etc) is increasing, but they have limited usefulness and do not address the most important clinical questions around this technology. Finally, the type and quality of information are a necessary condition to guarantee the usability of RWD to produce RWE. RWD/RWE could be an important source for HTA, but, as of today, there is much to be done before they can provide the main source of evidence for regulatory and/or reimbursement decision making. Details on the variables collected are often available through the case report form. Much more rarely information on missing values, outliers, and measurement errors are provided; information on completeness, for example, is reported almost only for registries. The provision of meta-data describing in more detail the content of data sources might improve the assessment of data integrity and quality. Existing scientific methods can help to address bias descending from noninterventional data, but their applicability is conditional on data quality.

### Strengths and Limitations

A key strength of this RWD mapping is that findings are the result of international collaboration and, therefore, we were able to use local networks and knowledge to explore RWD sources available in different countries and contexts in depth. Although this is a unique strength, it also brings with it some limitations, because the depth and scope of the mapping is subject to variability,

depending on the researchers' network. This may have played a role particularly in those countries where our authors were not based. We partly addressed this limitation with our systematic search on PubMed and through expert advice.

The other strength of this work is that we did not focus on specific RWD types, as it has often been done, especially for registries on hip and knee prostheses (e.g., Lübbecke et al. (22)). Moreover, we considered different medical devices and we were able to characterize the relevant types of RWD sources for each case not least through comparison.

Finally, it should be noted that the nature and the purpose of our mapping were not to identify and cover all existing sources. Rather, we focused hard on the availability and usability of RWD for HTA. To this end, we identified key sources (e.g., national registries and large observational studies) and we captured as many small-scale data sources as possible (e.g., multihospital databases). Full coverage is beyond the scope of this work and perhaps beyond the scope of any study, given that some sources are not accessible and not identifiable through publications, personal knowledge, network, or other means.

### Conclusions and Policy Recommendations

This mapping exercise and critical appraisal of RWD sources in Europe, though focused on three specific case studies, allows for general conclusions and policy recommendations beyond the selected medical technologies.

First, the development of standardized approaches for the design of RWD studies to make them more useful for HTA would be warranted. Stronger coordination at the EU level to leverage existing registries on devices currently scattered across different Member States would be highly beneficial for allowing the generation of comparable RWD across countries. These data would also help in establishing a solid evidence base for a more centralized HTA process, something presently under discussion in EU. Along these lines, the development of standardized meta-data describing data characteristics and their quality would represent a readily available tool for assessing whether the source is potentially suitable to HTA. The information presented in the extraction template of Supplementary Table 5, listing all RWD sources, represents a first attempt in this direction, as a possible list of elements to be described directly by data producers. Sound policy recommendations on how to improve the usability of RWD have already been proposed, for example for the use of registries in support of regulatory decision making: see, for example, the tools for assessing the usability of registries in support of regulatory decision making produced by the International Medical Device Regulators Forum (23).

Our study focuses on the content of the data sources and an assessment of their suitability to HTA. We are not interested in investigating the reasons for the setting up of specific registries/observational studies or on the process that leads to their creation. However, if (when) discussion at the EU level on the opportunities to anticipate the assessment of medical devices to earlier stages of their lifecycle (i.e., early dialogue/early HTA) develops further and the setting up of registries/observational studies becomes a regulatory requirement, the type of data collected will be of paramount importance to adequately inform policy decisions.

Second, more efforts should be made to leverage on RWD to produce comparative RWE. Comparative evidence can be derived, for example, from a control group taken from historical cohorts in

administrative databases or well-designed patient registries. Ideally, registries should be based on disease and should include more than one device. They should also routinely collect information on possible confounding factors and collect data on treatment patterns and resource use over time.

Finally, to further strengthen the use of RWE, jurisdictions need to develop a coordinated approach to the initiation, design, and analysis of RWD, working together with manufacturers. Such initiatives have been undertaken in various jurisdictions under different labels (coverage with evidence development, performance-based or risk-sharing agreements). They represent an important opportunity to fill in evidence gaps.

Our study sheds light on a series of challenges that RWD sources have if they are to be used for HTA purposes. If RWD are to be considered an important source of evidence on the economic and health impact of medical devices, there is a clear need to improve quality, quantity, and access to these data sources. This can only be achieved by coordinated and coherent actions across different stakeholders and jurisdictions.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S0266462321000301>.

**Acknowledgments.** The results reflect only the authors' views, and the EU is not responsible for any use that may be made of the information it contains.

**Funding.** This project received funding from the European Union's Horizon 2020 research and innovation program under grant agreement #779306 (COMED—Pushing the Boundaries of Cost and Outcome Analysis of Medical Technologies).

**Conflict of Interest.** HB reports grants from Celgene BV, outside the submitted work.

### References

1. Oortwijn W. Real-world evidence in the context of health technology assessment processes—from theory to action. HTAi Global Policy Forum. 2018.
2. Real-world data—Fact sheet [Internet] 2018 [updated April]. Available from: [https://ec.europa.eu/research/health/pdf/factsheets/real\\_world\\_data\\_factsheet.pdf](https://ec.europa.eu/research/health/pdf/factsheets/real_world_data_factsheet.pdf).
3. Schnell-Inderst P, Iglesias CP, Arvandi M, Ciani O, Matteucci Gothe R, Peters J, et al. A biased-adjusted evidence synthesis of RCT and observational data: The case of total hip replacement. *Health Econ*. 2017;26:46–69.
4. Schneeweiss S. Real-world evidence of treatment effects: The useful and the misleading. *Clin Pharmacol Ther*. 2019;106:43–4.
5. Franklin JM, Glynn RJ, Suissa S, Schneeweiss S. Emulation differences vs. biases when calibrating real-world evidence findings against randomized controlled trials. *Clin Pharmacol Ther*. 2020;107:735–7.
6. Franklin JM, Glynn RJ, Martin D, Schneeweiss S. Evaluating the use of nonrandomized real-world data analyses for regulatory decision making. *Clin Pharmacol Ther*. 2019;105:867–77.
7. Fleurence RL, Shuren J. Advances in the use of real-world evidence for medical devices: An update from the national evaluation system for health technology. *Clin Pharmacol Ther*. 2019;106:30.
8. Bowrin K, Briere J, Levy P, Millier A, Clay E, Toumi M. Cost-effectiveness analyses using real-world data: An overview of the literature. *J Med Econ*. 2019;22:545–53.
9. Tarricone R, Boscolo PR, Armeni P. What type of clinical evidence is needed to assess medical devices? *Eur Respir Rev*. 2016;25:259–65.
10. Murphy G. *Use of real-world evidence in single-drug assessments*. Ottawa, ON: CADTH; 2018.
11. US Department of Health and Human Services, US Food and Drug Administration. *Use of real-world evidence to support regulatory decision-making for Medical devices [cited 2021 March 30]*. 2021. Available



- from: <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf>.
12. Oortwijn W, Sampietro-Colom L, Trowman R. How to deal with the inevitable: Generating real-world data and using real-world evidence for HTA purposes—from theory to action. *Int J Technol Assess Health Care*. 2019;35:346–50.
  13. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health*. 2017;20:858–65.
  14. Leon MB, Piazza N, Nikolsky E, Blackstone EH, Cutlip DE, Kappetein AP, et al. Standardized endpoint definitions for transcatheter aortic valve implantation clinical trials: A consensus report from the Valve Academic Research Consortium. *J Am Coll Cardiol*. 2011;57:253–69.
  15. Kappetein AP, Head SJ, Génèreux P, Piazza N, Van Mieghem NM, Blackstone EH, et al. Updated standardized endpoint definitions for transcatheter aortic valve implantation: The Valve Academic Research Consortium-2 consensus document. *J Am Coll Cardiol*. 2012;60:1438–54.
  16. Pascual I, Carro A, Avanzas P, Hernández-Vaquero D, Díaz R, Rozado J, et al. Vascular approaches for transcatheter aortic valve implantation. *J Thorac Dis*. 2017;9:S478.
  17. Johnston SS, Chitnis AS, Gagne JJ, Ernst FR. *Medical Device Real-World Evidence for Beginners: A Primer*. Value & Outcomes Spotlight. 2019.
  18. Capodanno D. Transcatheter aortic valve implantation and mitral valve repair: Two trains, two speeds. *EuroIntervention*. 2017;12:1921.
  19. Gorst SL, Gargon E, Clarke M, Smith V, Williamson PR. Choosing important health outcomes for comparative effectiveness research: An updated review and identification of gaps. *PLoS One*. 2016;11:e0168403.
  20. Meregaglia M, Ciani O, Banks H, Salcher-Konrad M, Carney C, Jayawardana S, et al. A scoping review of core outcome sets and their “mapping” onto real-world data using prostate cancer as a case study. *BMC Med Res Methodol*. 2020;20:1–20.
  21. Khosravi A, Wendler O. TAVI 2018: From guidelines to practice. *Eur J Cardiol Pract*. 2018;15:29.
  22. Lübbecke A, Silman AJ, Barea C, Prieto-Alhambra D, Carr AJ. Mapping existing hip and knee replacement registries in Europe. *Health Policy*. 2018;122:548–57.
  23. International Medical Device Regulators Forum (IMDRF). *Tools for Assessing the Usability of Registries in Support of Regulatory Decision-Making [cited 2021 March 30]*. 2018. Available from: <http://www.imdrf.org/consultations/cons-registries-n46-pd1-170817.asp>.